COMPONENTS OF NEGRO-WHITE INCOME DIFFERENCES*

H. H. Winsborough and Peter Dickinson
Center for Demography and Ecology
University of Wisconsin

Introduction

This paper investigates black-white income differences for urban males in the labor force in 1960. Specifically, it addresses the question, "How much of the observed difference can be accounted for by differences in the educational level of the two populations?" In so doing, we have generalized components-of-a difference techniques somewhat.1/ That generalization is the most interesting methodological or statistical aspect of our paper.

We begin by discussing the method. Then we apply the technique to the analysis of Negro-white income differences. Finally, we will extend the generalization of our method to cover composit functions.

Components of a Difference

Suppose we have two functions on the real plane:

$$Y = f_1\ (X_1)$$
$$Y = f_2\ (X_2) \tag{1}$$

Let D_1 and D_2 denote the domains of f_1 and f_2 respectively, and R_1 and R_2 the co-domains. For any Y_1 an element of R_1, and Y_2 an element of R_2, there exists an $\underline{a}$ which is an element of D_1 and a $\underline{b}$, an element of D_2, such that $Y_1 = f_1^{1}(a)$ and $Y_2 = f_2\ (b)$. The difference $Y_1 - Y_2$ can be written as $f_1\ (a) - f_2\ (b)$.

Let:

$$\begin{aligned} \delta_1 &= f_2(a) - f_2(b) \\ \delta_2 &= f_1(b) - f_2(b) = \Delta f(b) \\ \delta_3 &= \Delta f(a) - \Delta f(b) \end{aligned} \tag{2}$$

Clearly,

$$\begin{aligned} \sum_{i=1}^{3} \delta_i &= [f_2(a) - f_2(b)] + [f_1(b) - f_2(b)] + \\ &\quad [f_1(a) - f_2(a) - f_1(b) + f_2(b)] \quad (3) \\ &= f_1(a) - f_2(b) = Y_1 - Y_2. \end{aligned}$$

We see then that the difference between two values, Y_1 and Y_2, can be expressed in terms of three additive components: a change in the argument of the function, a change in the function, and an interaction term, the result of a simultaneous change in both argument and function.

Application to Negro-white Income Differences

Suppose now that we let population one be urban white males in the labor force in 1960 while population two are comparably defined black males. Suppose, further, that we believe income is a linear function of education. In this example the Y_1 and X_1 of the above section are white income and education means respectively while Y_2 and X_2 are comparable Negro means. The function f_1 is the linear equation to predict income from education for whites while f_2 is the equation for blacks.

The component δ_1, then, is given by:

$$\begin{aligned} \delta_1 &= (\alpha_N + \beta_N \bar{X}_w) - (\alpha_N + \beta_N \bar{X}_N) = \\ &\quad \beta_N \bar{X}_w - \beta_N \bar{X}_N \end{aligned} \tag{4}$$

where the subscripts indicate the population of reference. This component can be interpreted as the gain in income which would result from an improvement of mean Negro education to the white level but with no change in the parameters of the Negro function.

The component δ_2 is given by:

$$\begin{aligned} \delta_2 &= (\alpha_w + \beta_w \bar{X}_N) - (\alpha_N + \beta_N \bar{X}_N) \\ &= (\alpha_w - \alpha_N) + (\beta_w - \beta_N) \bar{X}_N \end{aligned} \tag{5}$$

This component may be interpreted as the improvement which would result if Negro education were translated into income by the white rule but the level of Negro education were unchanged.

The component δ_3 is given by:

$$\begin{aligned} \delta_3 &= [(\alpha_w + \beta_w \bar{X}_w) - (\alpha_N + \beta_N \bar{X}_w)] - \\ &\quad [(\alpha_w + \beta_w \bar{X}_N) - (\alpha_N + \beta_N \bar{X}_N)] \\ &= (\beta_w - \beta_N)(\bar{X}_w - \bar{X}_N) \end{aligned} \tag{6}$$

In interpreting the preceding components we have used the rearranging of means or functions by some policy procedure as a kind of model to lend meaning to our components. If we continue using the policy model for interpretation, we might tell ourselves the following story about this component. Suppose the political or financial situation were such that it was possible to deal with only one aspect of the income discrimination problem at a time. Suppose a policy maker chose to improve the Negro mapping of education into income first. If that policy worked, the subsequent value of an improvement in Negro educational levels would no longer be simply the value of δ_1 but the sum of δ_1 and δ_3. Conversely, if we chose to improve Negro education levels first, the subsequent value of improving the Negro function would be the sum of δ_2 and δ_3.

Component three, then, is the increment (or decrement) in effect due to modifying both aspects of the situation simultaneously or in series over the effect of changing each singly.

The Data and Some Results

Let us now estimate these components using data from the one-in-a-thousand sample of the 1960 Census. We will estimate equations and components separately for three age groups, 25-34, 35-44, and 45-54. The dependent variable is total individual income scored at the mid-point of thousand dollar intervals. Educational attainment is measured by the highest grade of school completed scored as follows:

Years	Score
0-4	0
5-7	1
8	2
9-11	3
12	4
13-15	5
16	6
17 and above	7

Some preliminary analysis showed that these intervals yield the best approximation to linearity for the relationship between income and education. As we shall see presently, specification of a believable zero-point for educational attainment permits further analysis of component two -- the component for a change in functions. Since there is some reason to believe that functional literacy is generally obtained in the 5-7 interval, we have chosen to score 4 and fewer years of school as zero because attainment of functional literacy seems such a critical component of education.

Table 1 presents estimates of the six regression equations along with mean levels of education and income. All estimates are several times their standard deviation. Table 2 presents the components of the difference in mean income. Over all, whites make between \$2359 and \$3411 more money than do Negroes of the comparable ages and also have between 1.11 and 1.44 more units of education.

Table 1: Parameters for the regression of income on education by age group for Negro and for white urban males in the labor force,1960

	Age and Race					
	25-34		35-44		45-54	
Coefficient	White	Negro	White	Negro	White	Negro
Mean education	3.98	2.87	3.75	2.41	3.30	1.86
Mean income	4630	2270	5880	2670	5780	1821
bIE	572	219	1012	354	1128	295
aIE	2353	1641	2085	1817	2058	1821

Table 2: Components of Negro-white income difference by age

	Age		
Component	25-34	35-44	45-54
δ_1	243	474	425
δ_2	1725	1854	1786
δ_3	391	882	1200

Recall that the component δ_1 is the improvement in black income which would accrue if blacks had the same level of education as whites. From Table 2 we see that this improvement would be \$243 for the youngest age group, \$474 for the middle one and \$425 for the oldest. These quantities are between ten and fifteen percent of the total Negro-white income difference for the age group.

The component δ_2 is the improvement which would result if blacks had their own educational level but that education translated into income by the white equation. Estimating this component from our data we find the improvement in Negro income for the age groups would be \$1725 for the youngest, \$1854 for the middle and \$1786 for the oldest. These amounts range from 52 percent to 73 percent of the total difference.

Under the linear model it is possible to divide this component into two parts as shown in the second line of equation (5). One part has to do with differences in α and represents a fixed amount by which Negroes of <u>any</u> educational level are deprived of income in comparison to whites. The other part represents the difference in slope; the payoff of a unit of education. This separation of the component into two parts, however, depends on believing that one has set the zero-point of the education scale at the proper place. Had we set the zero-point of the educational scale at another place, the parts would divide differently but their sum would remain constant. With our assumption that 5 years of education represents functional literacy and all fewer years represent no education at all, we find change in the slope part of this component to be worth \$1013, \$1586, and \$1549 for the age group respectively. Leaving \$712, \$268, and \$237 attributable to intercept differences.

Returning to our main components, we find that δ_3, the interaction component has the value of \$391, \$882, and \$1200 for the age group. These amounts are about 17, 28, and 35 percent of the total difference.

From these data, then, it seems clear that δ_1, difference in educational level alone, is only a modest contribution to income differences. The component δ_2, differences in the equation mapping education into income, is by far the more important factor. Further, if one believes our setting of the zero-point for educational attainment, the villain can be narrowed down to differences in the return per unit of education. Perhaps this conclusion should have been apparent simply by an inspection of Table 1 where we find the differences in slopes ranging from \$833 to \$353. Even in this rather simple problem, we found some security of mind in working through the components to arrive at the conclusion. In more complex problems we have found the method invaluable. For that reason, we will conclude by extending our method somewhat.

Extensions of the Method

An important aspect of the components method presented above is that it holds not only for the class of continuous function but other kinds as well. Thus, the Kitagawa three component method is a special case of the method presented here.

In this regard it is worth noting that the Kitagawa two-component method has the virtue of a symmetry not obvious in our method. In the above example it has seemed natural to think of improving the Negro income to the white level. In that sense we have been using the white population as a standard. In another problem, it might not be clear which population should be standard. Thus, it is worth considering what would happen if we were to imagine reducing the white income level to that of blacks. Viewed from that perspective our components would be:

$$\begin{aligned} \delta_1' &= f_1(b) - f_1(a) \\ \delta_2' &= f_2(a) - f_1(a) = \Delta' f(a) \\ \delta_3' &= \Delta' f(b) - \Delta' f(a) \end{aligned} \tag{7}$$

Writing out δ_3' we see that

$$\begin{aligned} \delta_3' &= f_2(b) - f_1(b) - f_2(a) + f_1(a) = \\ &\quad \Delta f(a) - \Delta f(b) = \delta_3 \end{aligned} \tag{8}$$

Thus the interaction component is identical in the two approaches. Further, it is easy to show that:

$$\begin{aligned} -\ \delta_1' &= \delta_1 + \delta_3 \qquad \text{and} \\ -\ \delta_2' &= \delta_2 + \delta_3 \end{aligned} \tag{9}$$

Thus, in spite of the lack of symmetry in the three component model with changes in choice of the standard population, it is easy to move from one perspective to the other. Further, it is easy to show that the Kitagawa two-component method achieves its symmetry by simply dividing δ_3 among the other two components. In our notation her combined IJ effect is:

$$\frac{\delta_1 - \delta_1'}{2} = \frac{\delta_1 + \delta_1 + \delta_3}{2} = \delta_1 + \frac{\delta_3}{2} \tag{10}$$

while her residual IJ effect is

$$\frac{\delta_2 - \delta_2'}{2} \quad \frac{\delta_2 + \delta_2 + \delta_3}{2} \quad \delta_2 + \frac{\delta_3}{2} \tag{11}$$

The above methods are easily generalizable to composit functions. Suppose the Y-values are obtained from the composit of two functions,

$$\begin{aligned} Y &= f_1 \cdot g_1\ (x) \\ Y &= f_2 \cdot g_2\ (x) \end{aligned} \tag{12}$$

Letting $Y_1 = f_1g_1(a)$ and $Y_2 = f_2g_2(b)$ we can decompose the differences $Y_1 - Y_2$ as follows. Let

$$\begin{aligned} \delta_1 &= f_2g_2(a) - f_2g_2(b) \\ \delta_2 &= f_2g_2(b) - f_2g_2(b) \\ \delta_3 &= \Delta f g_2(b) \\ \delta_4 &= \Delta f \Delta g(b) \\ \delta_5 &= f_2\Delta g(a) - f_2\Delta g(b) \\ \delta_6 &= \Delta g(a) - \Delta g(b) \\ \delta_7 &= \Delta f \Delta g(a) - \Delta f \Delta g(b) \end{aligned} \tag{13}$$

It is easy to verify that:

$$\sum_{i=1}^{7} \delta_i = f_1g_1(a) - f_2g_2(b) \tag{14}$$

We have used these components to decompose income differences for a recursive system.

---

* The research reported here was supported by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity persuant to the provisions of the Economic Opportunity Act of 1964. The conclusions are the sole responsibility of the author. Certain data used in this analysis were derived by the authors from a computer tape file furnished under the joint project sponsored by the U. S. Bureau of the Census and the Population Council and containing selected 1960 Census information for a 0.1 percent sample of the population of the United States. Neither the Census Bureau nor the Population Council assumes any responsibility for the validity of any of the figures or interpretations of the figures presented herein or based on this material.

During the preparation of this report, Mr. Dickinson was supported as a trainee under NIH Training Grant #GM-01190 from the National Institute of General Medical Sciences.

1/ Evelyn M. Kitagawa, "Components of a Difference Between Two Rates," JASA, Vol. 50 (December 1955), pp. 1168-1194.